The R'enyi Entropy Function and the Large Deviation of Short Return Times

Nicolai Haydn, Sandro Vaienti

The Rényi Entropy Function and the Large Deviation of Short Return Times

Nicolai Haydn[*] Sandro Vaienti[†]

July 9, 2008

Abstract

We consider the Rényi entropy function for weakly ψ-mixing systems. The first main result proves existence and regularity properties. The second main result of the paper is to get the decay rate for the large deviation of the return time to cylinder sets. We show it to be exponential with a rate given by the Rényi entropy function. Finally we also obtain bounds for the free energy.

1 Introduction

The Rényi entropies [32] have been extensively studied in the eighties for their connections with various generalized spectra for dimensions of dynamically invariant sets, see for instance [23, 9, 18, 10, 22, 16, 29]. The commonly adopted definition generalizes the usual measure-theoretic entropy. Let T be a transformation on the measurable space Ω and μ a T-invariant probability measure on Ω. Assume Ω has a finite measurable partition $\mathcal{A}$ whose joins we denote by $\mathcal{A}^k = \bigvee_{j=0}^{k-1} T^{-j}\mathcal{A}$, $k = 1, 2, \dots$ (the elements of $\mathcal{A}^k$ are commonly referred to as n-cylinders). We assume $\mathcal{A}$ is generating, i.e. the elements of $\mathcal{A}^\infty$ are single points. For $t > 0$ we put

$$Z_n(t) = \sum_{A_n \in \mathcal{A}^n} \mu(A_n)^{1+t}$$

and define the *Rényi entropy function* $R_\mathcal{A}$ with respect to the partition $\mathcal{A}$ by

$$R_\mathcal{A}(t) = \liminf_{n\to\infty} \frac{1}{tn} |\log Z_n(t)|. \tag{1}$$ {Re}

Up to now the limit above has been proved to exist only in a few special situations: Bernoulli measures, Markov measures and more generally for Gibbs measures with Hölder continuous potentials ϕ: in fact in these cases the Rényi entropies can be expressed easily in terms of the topological pressure $P(\phi)$ of ϕ (see sect. 2.2 below) independently of the partition $\mathcal{A}$ (provided it is generating).

The *first* main result of our paper (Theorem 1) is to show the existence of the limit (1) for a wide class of measures (*dynamically weakly ψ-mixing measures*, see sect. 2.1). Moreover we prove that for $t \to 0^+$ the entropy $R_\mathcal{A}(t)$ converges to the metric-theoretic entropy $h(\mu)$, and that the function $tR_\mathcal{A}(t)$ is locally Lipschitz continuous.

In [34] Takens and Verbitsky suggested to define the Rényi entropy of order t of the measure-preserving transformation T as the function $\hat{R}(t) = \sup_\mathcal{A} R_\mathcal{A}(t)$ where the supremum is taken over all finite partitions $\mathcal{A}$ of Ω. This achieves that $\hat{R}(t)$ is a measure-theoretic invariant, but at the same

[*]Mathematics Department, USC, Los Angeles, 90089-1113. E-mail: <nhaydn@math.usc.edu>. This work was supported by a grant from the NSF (DMS-0301910) and by the CNRS unité FR2291 FRUMAN (#64636).

[†]Centre de Physique Théorique, UMR 6207, CNRS, Luminy Case 907, F-13288 Marseille Cedex 9, and Universities of Aix-Marseille I, II and Toulon-Var. Fédération de Recherche des Unités de Mathématiques de Marseille, France. E-mail: <vaienti@cpt.univ-mrs.fr>.

time becomes trivial since it was shown [34] that for ergodic measures μ the function $\hat{R}(t)$ is (for all $t > 0$) identically equal to the entropy $h(\mu)$. In order to "extract new information about the dynamics from the generalized entropies" [34], they introduced the *correlations entropies* by replacing cylinders by dynamical (Bowen) balls. The main application of correlation entropies was the complete characterization of the multifractal spectrum of local entropies for expansive homeomorphisms with specification [35] (see also [8, 31] for another approach). In fact in the latter case the correlation entropies coincide with the Rényi entropies $R_{\mathcal{A}}(t)$ computed with respect to any generating partition $\mathcal{A}$.

For the remainder of the paper we denote the Rényi entropy by $R(t)$ assuming that a given finite generating partition $\mathcal{A}$ has been chosen once and for all. The *second* main result of this paper (Theorem 4 and Corollary 5) uses the Rényi entropies to compute the large deviations of the first returns of cylinders A_n of length n. For this purpose let us introduce the return times function

$$\tau_A(x) = \min\left\{k \geq 1 : \ T^k x \in A\right\}$$

which is finite for μ-almost every $x \in A$ (Poincaré's theorem) and has expectation (on A) equal to 1 (Kac's theorem) when μ is ergodic.

For $n = 1, 2, \ldots$ let us put $\tau_n(x) = \min_{y \in A_n(x)} \tau_{A_n(x)}(y)$, where $A_n(x)$ denotes the n-cylinder that contains x. This quantity arose in several circumstances:
—Since it controls the short returns, it plays a crucial role to establish the asymptotic (exponential) distribution of the return times function $\tau_A(x)$ when the measure of the set A goes to zero [21, 3, 1, 2, 26, 25, 24].
—It has been used to define the *recurrence dimension* since it served as the gauge set function to construct a suitable Carathéodory measure [5, 30, 7].
—It has been related to the Algorithmic Information Content in [12].

The first result on the asymptotic behavior of $\tau_n(x)$ was proved in [33] (see also [6] for a different proof): for an ergodic measure μ of positive metric entropy $h(\mu)$, we have

$$\liminf_{n \to \infty} \frac{\tau_n(x)}{n} \geq 1 \qquad (2)$$ {SR}

for μ-almost every $x \in \Omega$. For systems which enjoy the specification property the preceding limit exists and is equal to 1 almost everywhere [33, 6]. The same result holds for a large class of maps on the interval with indifferent fixed points [20].

The situation changes considerably for systems with zero entropy. In general the limit (2) does not exist anymore and the values of the lim inf and lim sup depend upon the arithmetic properties of the map: see [27, 28, 14] for a careful investigation of Sturmian shifts and substitutive systems.

We will prove in Sect. 3 that the limit (2) exists almost surely and is equal to 1 even for weakly ψ-mixing measures. This leads immediately to the natural question of computing the large deviations for the process $\frac{\tau_n(x)}{n}$, namely to check the existence of the limit defining the lower deviation function

$$\lim_{n \to \infty} \frac{1}{n} \log \mu(x; \tau_n(x) \leq [\delta n]) \qquad (3)$$ {DF}

for $\delta \leq 1$. The case $\delta > 1$ is not interesting: it gives the value 0 to the above limit since $\tau_n(x) \leq n + \Delta$, where Δ is a constant independent of x and n, see Sect. 3.

The existence of the lower deviation function (3) was first established in [4] for classical ψ-mixing measures. These are special cases of the measures considered in this paper; they must satisfy the stronger mixing condition

$$\left| \frac{\mu(U \cap T^{-n-k}V)}{\mu(U)\mu(V)} - 1 \right| \leq \psi(k) \qquad (4)$$ {FM}

for all U in $\sigma(\mathcal{A}^n)$, for all n and all $V \in \sigma(\mathcal{A}^*)$ (the σ-algebra generated by $\mathcal{A}^n$), where we put $\mathcal{A}^* = \bigcup_{j=1}^{\infty} \mathcal{A}^j$ (compare this with the definition of weakly ψ-mixing measure introduced in subsection 2 below). The rate function $\psi(k)$, $k \geq 0$ must converges to zero, and, moreover, in order to achieve the existence of the limit (3) the additional assumption $\psi(0) < 1$ was required in [4]. This in particular implies (see Lemma 2.1 in [4]), that after having coded the elements of the initial finite partition (of cardinality $|\mathcal{A}| = M$, say) $\mathcal{A} = \bigcup_{i=1}^{M} A_i$ over the alphabet $G = \{1, 2, \ldots, M\}$ then for every string $\{i_0, \ldots, i_{n-1}\} \in G^n$, $n \geq 1$ the cylinder $A_{i_0} \cap T^{-1} A_{i_1} \cap \ldots \cap T^{(n-1)} A_{i_{n-1}}$ has positive measure, which essentially means that the grammar associated to the coding is complete. We will not anymore need this condition even for our larger class of weakly ψ-mixing measure. The key result in [4] was to relate the lower deviation function to the Rényi entropies for any ψ-mixing measure verifying the condition $\psi(0) < 1$, but in that paper the Rényi entropies were assumed to exist, since no general result was known.

It is well known that the deviation function could be computed as the Legendre transform of the free energy of the process, provided the free energy exists and is differentiable w.r.t. the parameter β (see (5) below). We show in Sect. 4 that this is not the case for our process: the free energy will be continuous but not differentiable at the point $\beta = -\gamma_\mu$, where γ_μ is the exponential decay rate of the measures of n-cylinders from Theorem 1. Even if the free energy is not differentiable, one can still derive an upper bound for the lower deviation function, which we will show to be consistent with the rigorous expression of the lower deviation function in terms of the Rényi entropies. It is interesting to note that the free energy function was also computed in [4], but the proof needed an additional assumption, namely the existence of a sequence of cylinders whose measure decays exponentially to zero with a rate which is exactly the constant γ_μ and whose first return is sublinear. We do not need anymore this hypothesis since we will prove the existence of such a sequence in full generality.

We gladly acknowledge numerous fruitful discussions with M Abadi about the Rényi entropy and its applications.

2 Rényi entropy function

{renyi}

2.1 Existence and regularity

We say the T-invariant probability measure μ on Ω is *weakly ψ-mixing* with respect to the (finite) partition $\mathcal{A}$ if there exist positive functions $\psi^-, \psi^+ : \mathbb{N} \to \mathbb{R}^+$, where $\psi^-(k) < 1 \ \forall\, k \geq \Delta_0$ for some Δ_0, so that

$$1 - \psi^-(k) \leq \frac{\mu(U \cap T^{-n-k} V)}{\mu(U)\mu(V)} \leq 1 + \psi^+(k)$$

for all U in $\sigma(\mathcal{A}^n)$, for all n and all $V \in \sigma(\mathcal{A}^*)$ (where $\mathcal{A}^* = \bigcup_{j=1}^{\infty} \mathcal{A}^j$). From now on we assume that the measure μ on Ω is a T-invariant non-atomic probability measure which is weakly ψ-mixing where the functions $1 - \psi^-, 1 + \psi^+$ are subexponential, which means $\limsup_{k\to\infty} \frac{1}{n} |\log(1 - \psi^-(k))| = 0$ and $\limsup_{k\to\infty} \frac{1}{n} \log(1+\psi^+(k)) = 0$. Lemma 3 shows that the measures of cylinder sets decay exponentially fast. Classical ψ-mixing measures correspond to the special case when $\psi^-(k) = \psi^+(k) = \psi(k)$ where $\psi(k) \searrow 0$ as $k \to \infty$ [11, 17, 15]. The classical ψ-mixing property implies in particular that μ cannot have any atoms.

Put $b_n = \max_{A_n \in \mathcal{A}^n} \mu(A_n)$ and let $\gamma_\mu = \liminf_n \frac{1}{n} |\log b_n|$ be *the exponential decay rate of the measures of n-cylinders*. We will now establish the following properties of the Rényi entropy:

{renyi.entro

Theorem 1 *Assume the (non-atomic) measure μ is weakly ψ-mixing and the functions $1 - \psi^-, 1 + \psi^+$ are subexponential. Then*
(I) *The limit $R(t) = \lim_{n\to\infty} \frac{1}{tn} |\log Z_n(t)|$ exists for $t > 0$. Convergence is uniform for t on compact subsets of $\mathbb{R}^+$.*
(II) *The function $W(t) = tR(t)$ is locally Lipschitz continuous.*
(III) $R(0) = \lim_{t\to 0^+} R(t) = h(\mu)$.
(IV) *$R(t)$ is monotonically decreasing on $(0, \infty)$ and $R(t) \to \gamma_\mu$ as $t \to \infty$, where $\gamma_\mu = \liminf_{n\to\infty} \frac{1}{n} |\log b_n|$ is positive.*

2.2 Examples

{examples}

(I) Bernoulli shift. If Ω is the full shift space over a finite alphabet $\{1, 2, \ldots, M\}$, σ the left shift transformation, the partition $\mathcal{A}$ is the collection of one-element cylinders and the invariant probability measure μ is given by a probability vector $\vec{p} = (p_1, p_2, \ldots, p_M)$ ($\sum_i p_i = 1$, $p_i > 0$), then $Z_n(t) = \left(\sum_i p_i^{1+t}\right)^n$ and the Rényi entropy is $R(t) = \frac{1}{t} \log \sum_i p_i^{1+t}$ for $t > 0$ and equal to the metric entropy $h_\mu = \sum_i p_i |\log p_i|$ for $t = 0$.

(II) Markov chains. Again Ω is the shift space over the alphabet $\{1, 2, \ldots, M\}$ and $\mathcal{A}$ is the usual partition of one-element cylinders. The invariant probability measure μ is now given by an $M \times M$ stochastic matrix P(we assume P is irreducible) and probability vector $\vec{p}$: $\vec{p}P = \vec{p}$ and $P1 = 1$. The cylinder set $U(x_1 \ldots x_n) \in \mathcal{A}^n$ which is given by the n-word $x_1 x_2 \ldots x_n$ then has the measure $\mu(x_1 \ldots x_n) = p_{x_1} P_{x_1 x_2} P_{x_2 x_3} \cdots P_{x_{n-1} x_n}$. Hence

$$Z_n(t) = \sum_{x_1 x_2 \ldots x_n} p_{x_1}^{1+t} P_{x_1 x_2}^{1+t} \cdots P_{x_{n-1} x_n}^{1+t}$$

where the sum is over all (admissable) n-words. The non-negative $M \times M$-matrix $P(t)$ whose entries are $P_{ij}(t) = P_{ij}^{1+t}$ has by the Perron-Frobenius theorem a single largest positive eigenvalue λ_t and a strictly positive (and normalised) left eigen vector $\vec{w}(t)$. (Note that λ_t is a continuous function of t and $\lambda_0 = 1$.) Thus ($p_i(t) = p_i^{1+t}$, $i = 1, \ldots M$)

$$\lambda_t^{-n} \vec{p}(t) P(t)^n \to (\vec{p} \cdot \vec{w}(t)) \vec{w}(t)$$

(exponentially fast) as $n \to \infty$. We thus obtain that $R(t) = \frac{1}{t} \log \lambda_t$ if t is positive and $R(0) = h_\mu = \sum_{ij} p_i P_{ij} |\log P_{ij}|$ if $t = 0$.

(III) Gibbs measures. [34, 10] If μ is a Gibbs measure for the potential function ϕ [13], then the Rényi entropy $R(t) = \frac{1}{t}((1+t)P(\phi) - P((1+t)\phi))$ (where P is the pressure function) is analytic for $t > 0$.

2.3 Proof of Theorem 1

Before we proof Theorem 1 we will need the following technical lemma about ψ-mixing measures. The notation $\psi^\pm$ means that ψ^+ applies when the left side inside the absolute value is positive and ψ^- applies when the argument inside the absolute value is negative.

{product.mix

Lemma 2 *Assume there are sets $B_j \in \sigma(\mathcal{A}^{n_j})$, $j = 1, 2, \ldots, k$ for some integers n_j. If μ is weakly ψ-mixing then*

$$\left| \mu\left(\bigcap_{j=1}^k T^{-N_j} B_j \right) - \prod_{j=1}^k \mu(B_j) \right| \le \left(\left(1 + \psi^\pm(\Delta)\right)^{k-1} - 1 \right) \prod_{j=1}^k \mu(B_j),$$

for all $\Delta \ge 0$, where $N_j = n_1 + n_2 + \cdots + n_{j-1} + (j-1)\Delta$ ($N_0 = 0$).

Proof. Put for $\ell = 1, 2, \ldots, k$:

$$D_\ell = \bigcap_{j=\ell}^k T^{-(N_j - N_\ell)} B_j.$$

In particular $\bigcap_{j=1}^k T^{-N_j} B_j = D_1$, $D_k = B_k$ and note that

$$D_\ell = B_\ell \cap T^{-n_\ell - \Delta} D_{\ell+1}.$$

By the mixing property $|\mu(D_\ell)-\mu(B_\ell)\mu(D_{\ell+1})|\le\psi^\pm(\Delta)\mu(B_\ell)\mu(D_{\ell+1})$ which repeatedly applied yields by the triangle inequality:

$$\begin{aligned}\left|\mu\left(\bigcap_{j=1}^k T^{-N_j}B_j\right)-\prod_{j=1}^k\mu(B_j)\right| &\le \psi^\pm(\Delta)\sum_{\ell=1}^{k-1}\mu\left(\bigcap_{j=1}^{\ell-1}T^{-N_j}B_j\right)\prod_{j=\ell}^{k-1}\mu(B_j)\\ &\le \left(\left(1+\psi^\pm(\Delta)\right)^{k-1}-1\right)\prod_{j=1}^k\mu(B_j).\end{aligned}$$

∎

{exponential

Lemma 3 *There exists a constant $\eta\in(0,1)$ so that $\mu(A_n)\le\eta^n$ for all $A_n\in\mathcal{A}^n$ and all n.*

Proof. Fix a $\Delta>0$ and m so that $b_m=\max_{A_m\in\mathcal{A}^m}\mu(A_m)\le\frac12(1+\psi^+(\Delta))^{-1}$ (note that $b_m\searrow 0$ as $m\to\infty$ since μ has no atoms). Then for any n (large) and $A_n\in\mathcal{A}^n$ one has $A_n\subset\bigcap_{j=0}^{k-1}T^{-km'}A_m(T^{jm'}A_n)$, where $k=[\frac{n}{m'}]$, $m'=m+\Delta$ and $A_m(T^{jm'}A_n)$ is the m-cylinder that contains $T^{jm'}A_n$ ($j\le k-1$). By Lemma 2

$$\mu(A_n)\le\mu\left(\bigcap_{j=0}^{k-1}T^{-km'}A_m(T^{jm'}A_n)\right)\le\left(1+\psi^+(\Delta)\right)^{k-1}\prod_{j=1}^k\mu(A_m(T^{jm'}A_n))\le\left(1+\psi^+(\Delta)\right)^k b_m^k\le 2^{-k}.$$

Hence $\eta\le 2^{-\frac{1}{m'}}$. ∎

Remark. The exponential decay of cylinders implies in particular that the metric entropy of a weakly ψ-mixing measure μ is positive. In fact $h(\mu)\ge|\log\eta|>0$.

Proof of (I). Let m and $\Delta\ge\Delta_0$ (the 'gap') be integers, put $m'=m+\Delta$ and let $n=km'-\Delta$ be a large integer. Put $\tilde{\mathcal{A}}^n=\bigvee_{j=0}^{k-1}T^{-jm'}\mathcal{A}^m$ define for some $\beta>1$

$$\mathcal{G}_n=\left\{A_n\in\mathcal{A}^n:\ \ \mu(A_n)\ge e^{-k\Delta^\beta}\mu(\tilde A_n)\right\},$$

where $\tilde A_n=\bigcap_{j=0}^{k-1}T^{-km'}A_m(T^{jm'}A_n)$. Then for every A_n one has

$$\mu\left(\bigcup_{A_n'\subset\tilde A_n,A_n'\in\mathcal{G}_n}A_n'\right)=\mu(\tilde A_n)-\mu\left(\bigcup_{A_n'\subset\tilde A_n,A_n'\notin\mathcal{G}_n}A_n'\right)\ge\left(1-|\mathcal{A}|^{k\Delta}e^{-k\Delta^\beta}\right)\mu(\tilde A_n)$$

as $\tilde A_n=\bigcup_{A_n'\subset\tilde A_n,A_n'\in\mathcal{A}^n}A_n'$ has k 'gaps' each of which is of length Δ. This implies that if $|\mathcal{A}|^\Delta e^{-\Delta^\beta}<1$ then for every $\tilde A_n\in\tilde{\mathcal{A}}^n$ there exists an $A_n'\subset\tilde A_n$, $A_n'\in\mathcal{A}^n$ which also belongs to $\mathcal{G}_n$. As $\Delta\ge\Delta_0$ we get

$$Z_n(t)=\sum_{A_n\in\mathcal{A}^n}\mu(A_n)^{1+t}\ge e^{-k\Delta^\beta(1+t)}\sum_{\tilde A_n\in\tilde{\mathcal{A}}^n}\mu(\tilde A_n)^{1+t}=e^{-k\Delta^\beta(1+t)}Z_m(t)^k\left(\left(1+\mathcal{O}(\psi^-(\Delta))\right)^{k-1}\right)^{1+t}$$

where we have used $\mu(\tilde A_n)=(1+\mathcal{O}(\psi^\pm(\Delta)))^{k-1}\prod_{j=0}^{k-1}\mu(A_m(T^{jm'}\tilde A_n))$ (mixing property). Hence we obtain

$$\begin{aligned}|\log Z_n(t)| &\le k|\log Z_m(t)|+k\Delta^\beta(1+t)+(1+t)\left|\log\left(1-\psi^-(\Delta)\right)^{k-1}\right|\\ &\le k|\log Z_m(t)|+\mathcal{O}(k\Delta^\beta(1+t))\end{aligned}$$

If we put $a_n = |\log Z_n(t)|$ then $a_n \le k a_m + ck\Delta^\beta$ and

$$\frac{a_{km'}}{km'} \le \frac{a_m}{m'} + c\frac{\Delta^\beta}{m'} = \frac{m}{m'}\frac{a_m}{m} + c\frac{\Delta^\beta}{m'}.$$

If we put $\Delta \sim m^\alpha$ so that $\alpha\beta < 1$ then $\limsup_n \frac{a_n}{n} \le \frac{m}{m+\Delta}\frac{a_m}{m} + \mathcal{O}(\frac{\Delta^\beta}{m})$ for all m. Hence $\limsup_n \frac{a_n}{n} \le \liminf_m \frac{a_m}{m}$.

We also have

$$Z_n(t) \le |\mathcal{A}|^{k\Delta} Z_m(t)^k (1+\mathcal{O}(\psi^+(\Delta)))^{(k-1)(1+t)}$$

which implies

$$|\log Z_n(t)| \ge k|\log Z_m(t)| + \mathcal{O}(k\Delta(1+t)).$$

This ensures uniform convergence for t in compact subsets of $\mathbb{R}^+$.

Proof of (II). For $t > 0$ let us put $H_n(t) = \sum_{A_n\in\mathcal{A}^n} \mu(A_n)^{1+t}|\log\mu(A_n)|$ (clearly $h(\mu) = \lim_{n\to\infty}\frac{1}{n}H_n(0)$ and $\frac{d}{dt}Z_n(t) = H_n(t)$). As above let $\tilde{\mathcal{A}}^n = \bigvee_{j=0}^{k-1} T^{-jm'}\mathcal{A}^m$ and, in order to cut k gaps of lengths Δ, put

$$\mathcal{G}_n = \left\{A_n \in \mathcal{A}^n : \ \mu(A_n) \ge e^{-k\Delta^\beta}\mu(\tilde{A}_n)\right\},$$

for some $\beta > 1$ where $\tilde{A}_n \in \tilde{\mathcal{A}}^n$ is so that $A_n \subset \tilde{A}_n$ and $n = km' - \Delta$ ($m' = m+\Delta$). The sum over $\mathcal{A}^n$ that defines Z_n is split into two parts: (i) over $\mathcal{G}_n$ and (ii) over the complement of $\mathcal{G}_n$.

(i) On the set $\mathcal{A}^n \setminus \mathcal{G}_n$ we have $\mu(A_n) \le e^{-k\Delta^\beta}\mu(\tilde{A}_n)$, where $A_n \in \mathcal{G}_n$, $A_n \subset \tilde{A}_n \in \tilde{\mathcal{A}}^n$. Choose $\gamma \in (1,\beta)$ and let $\mathcal{G}'_n = \left\{A'_n \in \mathcal{A}^n : \ \mu(A'_n) \ge e^{-k\Delta^\gamma}\mu(\tilde{A}_n)\right\}$. Then we get for all $A_n \notin \mathcal{G}_n$:

$$\mu(A_n) \le e^{-k\Delta^\beta}\mu(\tilde{A}_n) \le e^{-k\Delta^\beta}e^{-k\Delta^\gamma}\mu(A'_n),$$

where $A'_n \in \mathcal{G}'_n$ is so that $A'_n \subset \tilde{A}_n$ (such an A'_n exists since $|\mathcal{A}|^\Delta e^{-\Delta^\gamma} < 1$ for Δ large enough). Thus

$$\begin{aligned}
\sum_{A_n\notin\mathcal{G}_n} |\log\mu(A_n)|\,\mu(A_n)^{1+t} &\le e^{-k(1+t)\Delta^\beta}\sum_{A_n\notin\mathcal{G}_n}|\log\mu(A_n)|\,\mu(\tilde{A}_n)^{1+t}\\
&\le e^{-k(1+t)(\Delta^\beta-\Delta^\gamma)}\sum_{A'_n\in\mathcal{G}'_n}|\log\mu(A'_n)|\,\mu(A'_n)^{1+t}\\
&\le e^{-k(1+t)(\Delta^\beta-\Delta^\gamma)}H_n.
\end{aligned}$$

(ii) If $A_n \in \mathcal{G}_n$ then $\log\mu(A_n) = \log\mu(\tilde{A}_n) + \mathcal{O}(k\Delta^\beta)$ and we obtain

$$\begin{aligned}
H_n(t) &= \sum_{A_n\in\mathcal{A}^n}|\log\mu(A_n)|\mu(A_n)^{1+t}\\
&= \sum_{A_n\in\mathcal{G}_n}\left(|\log\mu(\tilde{A}_n)| + \mathcal{O}(k\Delta^\beta)\right)\mu(A_n)^{1+t} + \sum_{A_n\notin\mathcal{G}_n}|\log\mu(A_n)|\,\mu(A_n)^{1+t}\\
&= \sum_{A_n\in\mathcal{G}_n}|\log\mu(\tilde{A}_n)|\mu(A_n)^{1+t} + \mathcal{O}(k\Delta^\beta)Z_n + \mathcal{O}\left(e^{-k(1+t)(\Delta^\beta-\Delta^\gamma)}\right)H_n
\end{aligned}$$

(in the last step we used the estimate from part (i)).

The mixing property $\mu(\tilde{A}_n) = (1+\mathcal{O}(\psi^\pm(\Delta)))^{k-1}\prod_{j=0}^{k-1}\mu(A_m\circ T^{jm'})$ is applied to the principal term:

$$\sum_{A_n\in\mathcal{A}^n}|\log\mu(\tilde{A}_n)|\mu(A_n)^{1+t} = \sum_{j=0}^{k-1}X^j + \mathcal{O}(k(\psi^-(\Delta)+\psi^+(\Delta))),$$

where $X^j = \sum_{A_n\in\mathcal{A}^n} |\log\mu(A_m\circ T^{jm'})|\mu(A_n)^{1+t}$. To further examine X^j let us put

$$\tilde{\mathcal{A}}^n_j = \mathcal{A}^{jm'-\Delta}\vee T^{-jm'}\mathcal{A}^m\vee T^{-(j+1)m'-\Delta}\mathcal{A}^{n-(j+1)m'-\Delta}$$

where we opened up two gaps of lengths Δ ($\Delta\ge\Delta_0$), the first after j blocks and the second one after $j+1$ blocks ($j=0,\dots,k-1$) with the obvious modification if $j=0,k-1$ in which cases there is only a single gap. We now put

$$\mathcal{G}^j_n = \left\{A_n\in\mathcal{A}^n:\ \ \mu(A_n)\ge e^{-\Delta^\beta}\mu(\tilde{A}^j_n)\right\}$$

where $\tilde{A}^j_n\in\tilde{\mathcal{A}}^n_j$ is so that $A_n\subset\tilde{A}^j_n$. The sum in X^j over $\mathcal{A}^n$ is split into two parts: (a) over $\mathcal{G}^j_n$ and (b) over its complement $\mathcal{A}^n\setminus\mathcal{G}^j_n$.

(a) For the sum over $\mathcal{G}^j_n$ the mixing property

$$\mu(\tilde{A}^j_n) = (1+\mathcal{O}(\psi(\Delta)))\mu(A_{jm'-\Delta})\mu(A_m\circ T^{jm'})\mu(A_{n-(j+1)m'-\Delta}\circ T^{-(j+1)m'-\Delta})$$

for $\tilde{A}^j_n\in\mathcal{A}^n_j$ yields

$$\begin{aligned}\sum_{A_n\in\mathcal{G}^j_n}|\log\mu(A_m\circ T^{jm'})|\mu(A_n)^{1+t} &\in \left[e^{-(1+t)\Delta^\beta},|\mathcal{A}|^{2\Delta}\right]\sum_{\tilde{A}_n\in\tilde{\mathcal{A}}^n_j}|\log\mu(A_m\circ T^{jm'})|\mu(\tilde{A}^j_n)^{1+t}\\ &= \left[e^{-(1+t)\Delta^\beta},|\mathcal{A}|^{2\Delta}\right](1+\mathcal{O}(\psi^\pm(\Delta)))Z_{jm'-\Delta}H_mZ_{n-(j+1)m'-\Delta}.\end{aligned}$$

(b) For the sum over $\mathcal{A}^n\setminus\mathcal{G}^j_n$ we estimate as follows

$$\begin{aligned}\sum_{A_n\notin\mathcal{G}^j_n}|\log\mu(A_m\circ T^{jm'})|\mu(A_n)^{1+t} &\le |\mathcal{A}|^{2\Delta}e^{-(1+t)\Delta^\beta}\sum_{\tilde{A}^j_n\in\mathcal{A}^n_j}|\log\mu(A_m\circ T^{jm'})|\mu(\tilde{A}^j_n)^{1+t}\\ &\le |\mathcal{A}|^{2\Delta}e^{-(1+t)\Delta^\beta}(1+\mathcal{O}(\psi^\pm(\Delta)))^{1+t}Z_{jm'-\Delta}H_mZ_{n-(j+1)m'-\Delta}\end{aligned}$$

Similarly one shows that $Z_n\in\left[|\mathcal{A}|^{2\Delta}e^{-(1+t)\Delta^\beta},|\mathcal{A}|^{2\Delta}\right]Z_{jm'-\Delta}Z_mZ_{n-(j+1)m'-\Delta}$. Hence we get

$$H_n\in\left[\frac{1}{c_1},c_1\right]\sum_{j=0}^{k-1}\frac{Z_{jm'-\Delta}H_mZ_{n-(j+1)m'-\Delta}}{Z_n}+\mathcal{O}(k\Delta^\beta)$$

for some constant $c_1\approx 2|\mathcal{A}|^{2\Delta}e^{(1+t)\Delta^\gamma}$ and consequently

$$H_n\in\left[\frac{1}{c_1^2},c_1^2\right]kH_m+\mathcal{O}(k\Delta^\beta).$$

This implies that $\limsup_{n\to\infty}\frac{1}{n}H_n\le c_1^2\frac{1}{m}H_m$ and similarly $\liminf_{n\to\infty}\frac{1}{n}H_n\ge c_1^{-2}\frac{1}{m}H_m$. This implies that $c_2c_1^{-2}|s|\le W(t+s)-W(t)\le c_2c_1^2|s|$ for small s (e.g. $-t\le s\le 1$) for some positive constant c_2 (equal to $\frac{1}{m}H_m(t)$ for some m).

Proof of (III). With $H_n(t)=\sum_{A_n\in\mathcal{A}^n}\mu(A_n)^{1+t}|\log\mu(A_n)|$ as above we get

$$\begin{aligned}H_{n+m}(t) &= \sum_{A_{n+m}\in\mathcal{A}^{n+m}}\mu(A_{n+m})^{1+t}\left|\log\frac{\mu(A_{n+m})}{\mu(A_m)}+\log\mu(A_m)\right|\\ &= \sum_{A_{n+m}\in\mathcal{A}^{n+m}}\mu(A_{n+m})^{1+t}\left|\log\mu(A_m)\right|+\frac{1}{1+t}\sum_{A_{n+m}\in\mathcal{A}^{n+m}}\mu(A_{n+m})^{1+t}\left|\log\left(\frac{\mu(A_{n+m})}{\mu(A_m)}\right)^{1+t}\right|\\ &\le \sum_{A_m\in\mathcal{A}^m}\mu(A_m)^{1+t}\left|\log\mu(A_m)\right|+\frac{1}{1+t}\sum_{A_n\in\mathcal{A}^n}Z_m(t)\sum_{A_m\in\mathcal{A}^m}\frac{\mu(A_m)^{1+t}}{Z_m(t)}\phi\left(\left(\frac{\mu(A_{n+m})}{\mu(A_m)}\right)^{1+t}\right),\end{aligned}$$

where A_{n+m} stands for $A_m \cap T^{-m}A_n$ and $\phi(s) = -s\log s$ is concave on $(0,1)$ and increasing on $(0, \frac{1}{e})$. Thus

$$\begin{aligned}
H_{n+m}(t) &\leq H_m(t) + \frac{Z_m(t)}{1+t} \sum_{A_n \in \mathcal{A}^n} \phi\left(\sum_{A_m \in \mathcal{A}^m} \frac{\mu(A_{n+m})^{1+t}}{Z_m(t)} \right) \\
&\leq H_m(t) + \frac{Z_m(t)}{1+t} \sum_{A_n \in \mathcal{A}^n} \phi\left(\frac{\mu(A_n)^{1+t}}{Z_m(t)} \right),
\end{aligned}$$

provided $\frac{\mu(A_n)^{1+t}}{Z_m(t)} \leq \frac{1}{e}$ for every $A_n \in \mathcal{A}^n$. Hence

$$\begin{aligned}
H_{n+m}(t) &\leq H_m(t) + \frac{1}{1+t} \sum_{A_n \in \mathcal{A}^n} \mu(A_n)^{1+t} \left| \log \frac{\mu(A_n)^{1+t}}{Z_m(t)} \right| \\
&= H_m(t) + H_n(t) + \frac{1}{1+t} Z_n(t) |\log Z_m(t)|,
\end{aligned}$$

(as $Z_m \leq 1$). Now we apply this estimate repeatedly. In order to satisfy the condition $\frac{\mu(A_{jm})^{1+t}}{Z_m(t)} \leq \frac{1}{e}$ for every $A_{jm} \in \mathcal{A}^{jm}$, $j = 1, \dots, k$ let us note that the measure of the cylindersets goes to zero by Lemma 3. Hence for a given m we can find an integer J so that $\frac{\mu(A_{jm})^{1+t}}{Z_m(t)} \leq \frac{1}{e}$ for every $A_{jm} \in \mathcal{A}^{jm}$, and all $j > J$. Moreover since $W(0) = 0$ and $\frac{1}{n}\log Z_n(t)$ converge uniformly to $W(t)$ for $t \in (0,\delta)$ (for $\delta > 0$), we can let $\varepsilon > 0$ and choose $\delta > 0$ so that $|W(t)| < \frac{\varepsilon}{2}$ and N so that $|\frac{1}{n}\log Z_n(t) - W(t)| < \frac{\varepsilon}{2}$ for all $n \geq N$ and $t \in (0,\delta)$. Hence $1 \geq Z_n(t) \geq e^{-\varepsilon n}$ for $n \geq N, t \in (0,\delta)$. Assume $m > N$. Then we get almost subadditivity for the sequence $H_n(t)$:

$$H_{km}(t) = H_{Jm}(t) + (k-J)H_m(t) + \mathcal{O}(k\varepsilon m)$$

and consequently (as $k \to \infty$)

$$\lim_{n\to\infty} \frac{H_n(t)}{n} = \frac{H_m(t)}{m} + \mathcal{O}(\varepsilon)$$

for every $m > N$. Therefore if $t \in (0,\delta)$:

$$W(t) = \lim_{n\to\infty} \frac{\log Z_n(t)}{n} = \lim_{n\to\infty} \frac{1}{n} \int_0^t H_n(s)\,ds = \frac{1}{m} \int_0^t H_m(s)\,ds + \mathcal{O}(\varepsilon t)$$

and consequently

$$R(0) = W'(0) = \lim_{t\to 0^+} \frac{1}{tm} \int_0^t H_m(s)\,ds + \mathcal{O}(\varepsilon) = \frac{1}{m} H_m(0)\,ds + \mathcal{O}(\varepsilon).$$

Since $\varepsilon > 0$ was arbitrary we get that $R(0) = \lim_{m\to\infty} \frac{1}{m} H_m(0)$ (we need that $m > N_\varepsilon$, where $N_\varepsilon \to \infty$ as $\varepsilon \to 0^+$).

Proof of (IV). The fact that R is decreasing was noted in e.g. [34, 9]. Since

$$b_n^{1+t} \leq Z_n(t) \leq \sum_{A_n \in \mathcal{A}^n} \mu(A_n) b_n^t = b_n^t,$$

we obtain that $\frac{1}{n}|\log b_n| \leq R(t) \leq \frac{1+t}{t}\frac{1}{n}|\log b_n|$ for all n (this estimate is true universally, independent of mixing properties). Hence $\gamma_\mu \leq R(t) \leq \frac{1+t}{t}\gamma_\mu$ for all $t > 0$ where γ_μ is strictly positive since, by Lemma 3, $\gamma_\mu \geq |\log \eta| > 0$. ∎

As Lemma 3 shows the measure of cylinder sets always decays exponentially fast for weakly ψ-mixing measures. Clearly, if the measure of cylinder sets decays subexponentially (i.e. $\gamma_\mu = 0$) then the Rényi entropy $R(t)$ is identically zero on $(0, \infty)$.

3 Short return times

In the introduction we recalled that for every ergodic measure μ with positive entropy $\liminf_{n\to\infty}\frac{1}{n}\tau_n(x)\geq 1$ almost everywhere. Since a weakly ψ-mixing measure μ has positive entropy (see the remark following Lemma 3), we obtain $\liminf_{n\to\infty}\frac{1}{n}\tau_n(x)\geq 1$ for μ-almost every $x\in\Omega$. In order to get the upper bound let $x\in\Omega$, let us note that by the weak ψ-mixing property

$$\frac{\mu(A_n(x)\cap T^{-n-\Delta}A_n(x))}{\mu(A_n(x))^2}\geq 1-\psi^-(\Delta)>0$$

for $\Delta\geq\Delta_0$. This implies $\tau_n(x)\leq n+\Delta$ and since $\Delta\geq\Delta_0$ is fixed we obtain that $\limsup_{n\to\infty}\frac{1}{n}\tau_n(x)\leq 1$ for every $x\in\Omega$. Hence

$$\lim_{n\to\infty}\frac{1}{n}\tau_n(x)=1$$

almost everywhere. In this section we are concerned about the large deviations of the process τ_n, namely we are interested in the asymptotic behavior of the distributions

$$\mathbb{P}(\tau_n\leq[\delta n])=\mu(\{x:\tau_n(x)\leq[\delta n]\}).$$

Since $\tau_n(x)$ is obviously constant for all the points in the same cylinder around x, we could replace the set $\{x:\tau_n(x)\leq[\delta n]\}$ with the following one

$$\mathcal{C}_n(\delta)=\{A_n\in\mathcal{A}^n:\ \ \tau_n(A_n)\leq[\delta n]\}$$

which measures the probability of points to have very short returns and where $\tau_n(A_n)=\min\{k\geq 1: A_n\cap T^{-k}A_n\neq\emptyset\}=\tau_n(x)$, $x\in A_n$. In order to analyze its size let us, following [4], define the sets

$$B_n(j)=\left\{A_n\in\mathcal{A}^n:\frac{j}{\tau_n(A_n)}\in\mathbb{N}\right\},$$

where $n\in\mathbb{N}, j=1,\dots,n$. Clearly $B_n(j)\in\sigma(\mathcal{A}^n)$ for all j and if we look at the symbolic representation of the n-cylinders in $B_n(j)$ we note that there are two cases, namely:
(i) If $j\leq\frac{n}{2}$ and x is a point in $B_n(j)$ then the first n symbols of points in it are

$$(x_1x_2\dots x_j)^{n'}x_1x_2\dots x_r$$

where $n'=[n/j]$ and $r=n-j[n/j]$ $(r<j)$.
(ii) If $j>\frac{n}{2}$ and A_n is an n-cylinder in $B_n(j)$ then the first n symbols of points in it are

$$x_1x_2\dots x_{n-j}x_1x_2\dots x_{2j-n}x_1x_2\dots x_{n-j}$$

where the (remainder) middle portion is of length $n-2(n-j)=2j-n$.

Let us put

$$\mathcal{S}_n(\lambda)=\{A_n\in\mathcal{A}^n:\ \ \tau_n(A_n)=[n\lambda]\}\,.$$

The purpose of this section is to determine the decay rate of the measure of the set $\mathcal{S}_n(\lambda)$ as n goes to infinity. As λ varies over the unit interval we obtain the short recurrence spectrum for the measure μ. Let us note that for every n we have that $\mathcal{C}_n(\delta)=\bigcap_{\lambda<\delta}\mathcal{S}_n(\lambda)$

For $\lambda\in(0,1]$ we define the function

$$M(\lambda)=(1-\lambda\ell)\left(W(\ell)-W(\ell-1)\right)+\delta W(\ell-1)$$

where $d=\left[\frac{1}{\delta}\right]$ $(1-\delta d$ linearly interpolates between the values $\frac{1}{k+1}$ and 0 on the interval $(\frac{1}{k+1},\frac{1}{k}))$. The function $M(\lambda)$ is continuous on $(0,1)$, piecewise affine on the intevals $(\frac{1}{1+k},\frac{1}{k})$ and assumes the values $M(\frac{1}{k})=\frac{1}{k}W(k-1)$, $k=1,2,\dots$ (in particular $M(1)=0$). The function $M(\lambda)$ interpolated $\hat{M}(\lambda)=\frac{\lambda}{1+\lambda}W(\frac{1}{\lambda})$ between the values at the points $\lambda=\frac{1}{k}$ for $k=1,2,\dots$. Changing coordinates to

$t=\frac{1+\lambda}{\lambda}$ we get $\hat{M}(\lambda)=\frac{1}{t}W(t-1)$. This function is increasing for $t>1$ as can be seen from the derivatives of the approximating functions. To wit

$$\frac{d}{dt}\frac{1}{tn}|\log Z_n(t-1)|=\frac{1}{t^2Z_n(t-1)}\sum_{A_n\in\mathcal{A}^n}\mu(A_n)^t\left|\log\frac{\mu(A_n)^t}{Z_n(t-1)}\right|,$$

which is positive for every n. Since $\lim_{n\to\infty}\frac{1}{tn}|\log Z_n(t-1)|=\frac{1}{t}W(t-1)$ we conclude that $\frac{1}{t}W(t-1)$ is increasing on $(1,\infty)$. Hence $M(\lambda)$ is decreasing on $(0,1)$. We now prove our main result for the density of short returns.

{main.result

Theorem 4

$$\limsup_{n\to\infty}\frac{1}{n}|\log\mu(\mathcal{S}_n(\lambda))|=M(\lambda)$$

The lower bound was proven in [4]. It remains to prove the upper bound. In [4] the bound was proven under the assumption that $\psi(0)$ be less than 1 which is essentially only satisfied for Bernoulli measures. Here we obtain the lower bound for all weakly ψ-mixing measures. Theorem 4 gives rise to the following corollary.

{C.limit}

Corollary 5

$$\limsup_{n\to\infty}\frac{1}{n}|\log\mu(\mathcal{C}_n(\delta))|=M(\delta)$$

Proof. Clearly $\mathcal{C}_n(\delta)\subset\bigcup_{j=1}^{[\delta n]}B_n(j)$ which implies that $\mathcal{C}_n(\delta)\subset\bigcup_{0<\lambda\le\delta}\mathcal{S}_n(\lambda)$. The union in fact consists of no more than n distinct sets. Hence

$$\mu(\mathcal{C}_n(\delta))\le n\max_{0<\lambda\le\delta}\mu(\mathcal{S}_n(\lambda))$$

which implies that $\limsup_{n\to\infty}\frac{1}{n}|\log\mu(\mathcal{S}_n(\lambda))|\le\min_{0<\lambda\le\delta}M(\lambda)$. The upper bound follows from the fact that $\mathcal{S}_n(\lambda)\subset\mathcal{C}_n(\delta)$ for every $\lambda\le\delta$. The statement now follows because M is monotonically decreasing on $(0,1)$. ∎

{upper.S.bou

Proposition 6

$$\limsup_{n\to\infty}\frac{1}{n}|\log\mu(\mathcal{S}_n(\lambda))|\le M(\lambda)$$

Let us first prove the following inequality which by itself is of some interest.

{lower.Z.bou

Lemma 7 *Let $\gamma\in(0,1)$. Then for all $\lambda\in(0,1)$ and all large enough n:*

$$\mu(B_n(j))\ge e^{\mathcal{O}(n^\gamma)}Z_r(w)Z_{j-r}(w-1),$$

where $j=[\lambda n]$ and $n=wj+r$, $0\le r<j$, $w=[n/j]$.

Proof. We do the two cases (A) $\lambda\in(0,\frac{1}{2}]$ and (B) $\lambda\in(\frac{1}{2},1)$ separately.

(A) Let us first deal with the case $0<\lambda\le\frac{1}{2}$. Put $j=[\lambda n]$ and $w=\left[\frac{n}{j}\right]$. Then $n=wj+r$ where $r<j$ ($r=0$ if $\lambda n\in\mathbb{N}$ and $1/\lambda\in\mathbb{N}$). For an n-cylinder $A_n\subset B_n(j)$ one has the decomposition

$$A_n=\left(\bigcap_{k=0}^{w-1}T^{-jk}A_j(A_n)\right)\cap T^{-wj}A_r(A_n),$$

where $w=\left[\frac{n}{j}\right]$ and $A_j(A_n)$ is the j-cylinder that contains the n-cylinder A_n etc. Let $\Delta\ge\Delta_0$ be so that $\Delta<r,j-r$ and put

$$\tilde{A}_n=\left(\bigcap_{k=0}^{w}T^{-jk}A_{r-\Delta}(A_n)\right)\cap\left(\bigcap_{k=0}^{w}T^{-jk-r}A_{j-r-\Delta}(A_n)\right).$$

Here we opened up gaps of lengths Δ (i) at each occurrence of period after length j and (ii) then cut each period of length j into two pieces of lengths r and $r-j$. Since $A_n \subset \tilde{A}_n$ clearly $\mu(\tilde{A}_n) \geq \mu(A_n)$ and in order to get a comparison in the opposite direction, let $\beta > 1$ and put

$$\mathcal{G}_{n,j} = \left\{ A_n \in \mathcal{A}^n :\ A_n \subset B_n(j),\ \ \mu(A_n) \geq e^{-2w\Delta^\beta} \mu(\tilde{A}_n) \right\},$$

for the 'good' n-cylinders in $B_n(j)$ whose measures are comparable to the measure of $\tilde{A}_n$. Put $G_{n,j} = \bigcup_{A_n \in \mathcal{G}_{n,j}} A_n$. Then for every $A_n \subset B_n(j)$ one has

$$\mu\left(\bigcup_{A'_n \subset \tilde{A}_n \cap B_n(j), A'_n \in \mathcal{A}^n} A'_n \right) \geq \left(1 - |\mathcal{A}|^{2w\Delta} e^{-2w\Delta^\beta}\right) \mu(\tilde{A}_n \cap B_n(j))$$

as $\tilde{A}_n \cap B_n(j) = \bigcup_{A'_n \subset \tilde{A}_n, A'_n \in \mathcal{G}_{n,j}} A'_n$. This implies that if $|\mathcal{A}|^{2w\Delta} e^{-2w\Delta^\beta} < 1$ then $\tilde{A}_n \cap B_n(j) \neq \emptyset$ if and only if there exists an $A'_n \subset \tilde{A}_n$, $A'_n \in \mathcal{A}^n$, which also belongs to $\mathcal{G}_{n,j}$. Hence

$$\mu(B_n(j)) \geq \mu(G_{n,j}) \geq e^{-2w\Delta^\beta} \sum_{\tilde{A}_n} \mu(\tilde{A}_n),$$

where the sum is over all $\tilde{A}_n$ for which there is an $A_n \subset B_n(j)$. Since all $A_n \subset B_n(j)$ are of the form $(x_1 \dots x_j)^w x_1 \dots x_r$ where $x_1 \dots x_j$ runs through all possible periodic words of lengths j, we get

$$\begin{aligned} \sum_{\tilde{A}_n} \mu(\tilde{A}_n) &= \left(1 + \mathcal{O}(\psi^{\pm}(\Delta))\right)^{2w+1} \sum_{x_1 \dots x_{r-\Delta}} \sum_{x_{r+1} \dots x_{j-\Delta}} \mu(A_{r-\Delta}(x_1 \dots x_{r-\Delta}))^{w+1} \mu(A_{j-r-\Delta}(x_{r+1} \dots x_{j-\Delta}))^w \\ &= \left(1 + \mathcal{O}(\psi^{\pm}(\Delta))\right)^{2w+1} Z_{r-\Delta}(w) Z_{j-r-\Delta}(w-1), \end{aligned}$$

where the sum is over all $(r-\Delta)$-words $x_1 \dots x_{r-\Delta}$ and all $(j-n-\Delta)$-words $x_{r+1} \dots x_{j-\Delta}$, where $Z_m(k) = \sum_{A_m \in \mathcal{A}^m} \mu(A_m)^{k+1}$. Hence

$$\mu(B_n(j)) \geq c_1 e^{-2w\Delta^\beta} Z_{r-\Delta}(w) Z_{j-r-\Delta}(w-1)$$

for some $c_1 > 0$. We have to choose $\Delta \geq \Delta_0$ and need to have $|\mathcal{A}|^\Delta e^{-\Delta^\beta} < 1$. This requires β to be bigger than 1.

Next we compare $Z_{r-\Delta}(w)$ to $Z_r(w)$ as follows:

$$Z_{r-\Delta}(w) = \sum_{x_1 \dots x_{r-\Delta}} \mu(A_{r-\Delta}(x_1 \dots x_{r-\Delta}))^{w+1} \geq \frac{1}{|\mathcal{A}|^\Delta} \sum_{x_1 \dots x_r} \mu(A_{r-\Delta}(x_1 \dots x_r))^{w+1}$$

as $\mu(A_r(x_1 \dots x_r)) \leq \mu(A_{r-\Delta}(x_1 \dots x_{r-\Delta}))$ and $\#\{A_r : A_r \subset A_{r-\Delta}\} \leq |\mathcal{A}|^\Delta$. Hence $Z_{r-\Delta}(w) \geq |\mathcal{A}|^{-\Delta} Z_r(w)$ and similarly $Z_{j-r-\Delta}(w-1) \geq |\mathcal{A}|^{-\Delta} Z_{j-r}(w-1)$. This implies

$$\mu(B_n(j)) \geq c_1 |\mathcal{A}|^{-2\Delta} e^{-2w\Delta^\beta} Z_r(w) Z_{j-r}(w-1) \geq c_1 e^{-c_2 \Delta^{\alpha\beta}} Z_r(w) Z_{j-r}(w-1)$$

if we choose $\Delta = [j^\alpha]$ for some $\alpha \in (0,1)$. If α is small enough then $\gamma \geq \alpha\beta$.

(B) The case when $\lambda \in (\frac{1}{2}, 1)$. Again we put $j = [n\lambda]$ and $n = j + r$ (note $[1/\lambda] = 1$). If $A_n \subset B_n(j)$ is an n-cylinder then $A_n = A_r(A_n) \cap T^{-j} A_r(A_n) \cap T^{-r} A_{n-2r}(T^r A_n)$, where $n - 2r \geq 0$. Let $\Delta \geq \Delta_0$ (not too large) and define as above

$$\tilde{A}_n = A_{r-\Delta}(A_n) \cap T^{-j} A_{r-\Delta}(A_n) \cap T^{-r} A_{n-2r-\Delta}(T^r A_n)$$

(of $n - 2r - \Delta > 0$, otherwise we just put $\tilde{A}_n = A_{r-\Delta}(A_n) \cap T^{-j} A_{r-\Delta}(A_n)$). For $\beta > 1$ we introduce as before the 'good set'

$$\mathcal{G}_{n,j} = \left\{ A_n \in \mathcal{A}^n :\ \ A_n \subset B_n(j),\ \mu(A_n) \geq e^{-\Delta^\beta} \mu(\tilde{A}_n) \right\}.$$

If $|\mathcal{A}|^{2\Delta}e^{-\Delta^\beta}<1$ then for every $\tilde{A}_n$ (of the form given above) there exists an $A_n\in\mathcal{G}_{n,j}$ so that $A_n\subset\tilde{A}_n$ and therefore

$$\mu(B_n(j))\geq\mu(G_{n,j})\geq e^{-\Delta^\beta}\sum_{\tilde{A}_n}\mu(\tilde{A}_n),$$

where the sum is over all $\tilde{A}_n=A_{r-\Delta}(x_1\dots x_{r-\Delta})\cap T^{-j}A_{r-\Delta}(x_1\dots x_{r-\Delta})\cap T^{-r}A_{r-\Delta}(x_{r+1}\dots x_{n-j})$ (in the case when $n-2r-\Delta>0$) and $x_1\dots x_{r-\Delta}$, $x_{r+1}\dots x_{n-j}$ are arbitrary words. Hence

$$\begin{aligned}\mu(B_n(j)) &\geq \left(1+\mathcal{O}(\psi^-(\Delta))\right)^2e^{-\Delta^\beta}\sum_{x_1\dots x_{r-\Delta}}\mu(A_{r-\Delta}(x_1\dots x_{r-\Delta}))^2\sum_{x_{r+1}\dots x_{j-\Delta}}\mu(A_{j-r-\Delta}(x_{r+1}\dots x_{j-\Delta}))\\ &= \left(1+\mathcal{O}(\psi^-(\Delta))\right)^2e^{-\Delta^\beta}Z_{r-\Delta}(1)Z_{j-r-\Delta}(0)\\ &= \left(1+\mathcal{O}(\psi^-(\Delta))\right)^2|\mathcal{A}|^{-2\Delta}e^{-\Delta^\beta}Z_r(1)Z_{j-r}(0)\end{aligned}$$

where in the last line we used the comparison from the end of part (A). Again we choose $\Delta=[j^\alpha]$ where $\alpha\in(0,1)$ can be chosen small enough so that $\gamma\geq\alpha\beta$. ∎

Proof of Proposition 6. Obviously $\mu(\mathcal{S}_n(\lambda))\geq B_n(j)$ and therefore by Lemma 7

$$\frac{\log\mu(\mathcal{S}_n(\lambda))}{n}\geq-\frac{\mathcal{O}(n^\gamma)}{n}+\frac{1}{n}\log Z_r(w)+\frac{1}{n}\log Z_{j-r}(w-1)$$

where $\gamma<1$ can be chosen arbitrarily. As $n\to\infty$ the first term goes to zero. Thus

$$\liminf_{n\to\infty}\frac{1}{n}\log\mu(\mathcal{S}_n(\lambda))\geq\liminf_{n\to\infty}\frac{1}{n}\log Z_r(w)+\liminf_{n\to\infty}\frac{1}{n}\log Z_{j-r}(w-1)$$

Now notice that (as $n=w[\lambda n]+r$)

$$\frac{1}{n}\log Z_r(w)=\frac{r}{n}\frac{1}{r}\log Z_r(w)\to(1-\lambda\ell)\,W(\ell)$$

since $\frac{r}{n}=\frac{n-[\lambda n]w}{n}=1-\frac{[\lambda n]}{n}w\to1-\lambda\ell$ and $w\to\ell$ as $n\to\infty$ and W is continuous by Theorem 1. Similarly

$$\frac{1}{n}\log Z_{j-r}(w)=\frac{j-r}{n}\frac{1}{j-r}\log Z_{j-r}(w)\to(\lambda\,(1+\ell)-1)\,W(\ell-1)$$

since $\frac{j-r}{n}=\frac{[n\lambda]-(n-[\lambda n])w}{n}=\frac{[n\lambda]}{n}(1+w)-1\to\lambda\,(1+\ell)-1$. This implies the statement of the proposition. ∎

{lower.S.bou

Lemma 8

$$\limsup_{n\to\infty}\frac{1}{n}|\log\mu(\mathcal{S}_n(\lambda))|\geq M(\lambda)$$

Proof. Again we do the two cases (A) $\lambda\in(0,\frac12]$ and (B) $\lambda\in(\frac12,1)$ separately.
(A) $0<\lambda\leq\frac12$: We decompose as above $n=wj+r$, where $j=[\lambda n]$ and $w=[n/j]$, $0\leq r<j$. Since all $A_n\subset B_n(j)$ are of the form $(x_1\dots x_j)^wx_1\dots x_r$ where $x_1\dots x_j$ runs through all possible periodic words of lengths j, we get (summing over such A_n)

$$\sum_{A_n}\mu(A_n)\leq\left(1+\psi^+(0)\right)^{2w+1}\sum_{x_1\dots x_r}\sum_{x_{r+1}\dots x_j}\mu(A_r(x_1\dots x_r))^{w+1}\mu(A_{j-r}(x_{r+1}\dots x_j))^w,$$

where the sum is over all r-words $x_1\dots x_r$ and all $(j-n)$-words $x_{r+1}\dots x_j$. Hence

$$\mu(\mathcal{S}_n(\lambda))\leq\left(1+\psi^+(0)\right)^{2w+1}Z_r(w)Z_{j-r}(w-1)$$

and therefore as in the proof of Proposition 6

$$\lim_{n\to\infty}\frac{1}{n}|\log\mu(\mathcal{S}_n(\lambda)|\geq\lim_{n\to\infty}\frac{1}{n}|\log Z_r(w)|+\lim_{n\to\infty}\frac{1}{n}|\log Z_r(w-1)|=M(\lambda).$$

(B) $\lambda \in (\frac{1}{2}, 1)$: Again we put $j = [n\lambda]$ and $n = j + r$, $0 \le r < j$ (as $[1/\lambda] = 1$). If $A_n \subset B_n(j)$ is an n-cylinder then $A_n = A_r(A_n) \cap T^{-j} A_r(A_n) \cap T^{-r} A_{n-2r}(T^r A_n)$, where $n - 2r \ge 0$. Hence

$$\mu(B_n(j)) \le \left(1 + \psi^+(0)\right)^2 \sum_{x_1 \dots x_r} \mu(A_r(x_1 \dots x_r))^2 \sum_{x_{r+1} \dots x_j} \mu(A_{j-r}(x_{r+1} \dots x_j))$$

for arbitrary words $x_1 \dots x_r$, $x_{r+1} \dots x_{n-j}$. This implies $\mu(\mathcal{S}_n(\lambda)) \le c_1 Z_r(1) Z_{j-r}(0)$ ($c_1 > 0$) and $\lim_{n\to\infty} \frac{1}{n} |\log \mu(\mathcal{S}_n(\lambda)| \ge \lim_{n\to\infty} \frac{1}{n} |\log Z_r(1)| = M(\lambda)$. ∎

Proof of Theorem 4. The theorem now follows from Proposition 6 and Lemma 8. ∎

Apart from the exact limiting behaviour we get from Theorem 4 we can also proof the following simpler bounds.

Lemma 9

$$\liminf_{n\to\infty} \frac{1}{n} |\log \mu(\mathcal{C}_n(\delta))| \le h_\mu(1 - \delta)$$

for all $\delta \in (0, 1)$.

Proof. As before $j = [\delta n]$ and for an n-cylinder A_n in $B_n(j)$ we put $\tilde{A}_n = A_{r-\Delta}(A_n) \cap T^{-r} A_{n-r}(T^r A_n)$ for a gap of length Δ on the segment $[r - \Delta + 1, r]$. As before we let $\beta > 1$,

$$\mathcal{G} = \left\{A_n \in \mathcal{A}^n : \;\; A_n \subset B_n(j),\; \mu(A_n) \ge e^{-\Delta^\beta} \mu(\tilde{A}_n)\right\}$$

and observe that if $|\mathcal{A}|^\Delta e^{-\Delta^\beta} < 1$ then

$$\tilde{A}_n \cap B_n(j) \ne \emptyset \Longleftrightarrow \exists A'_n \in \mathcal{G},\; A'_n \subset \tilde{A}_n$$

Hence ($G = \bigcup_{A_n \in \mathcal{G}} A_n$) if $\Delta \ge \Delta_0$ then

$$\begin{aligned}
\mu(B_n(j)) &\ge \mu(G) \\
&\ge e^{-\Delta^\beta} \sum_{\tilde{A}_n} \mu(\tilde{A}_n) \\
&\ge \left((1 - \psi^-(\Delta))\right) e^{-\Delta^\beta} \sum_{x_1 \dots x_{j-\Delta}} \mu(A_{j-\Delta}(x_1 \dots x_{j-\Delta})) \sum_{x_{j-\Delta+1} \dots x_j} \mu(A_{n-j}(T^j (x_1 \dots x_j)^\infty))
\end{aligned}$$

as $\tilde{A}_n = A_{j-\Delta}(x_1 \dots x_{j-\Delta}) \cap A_{n-j}((x_1 \dots x_j)^\infty)$. By the Shannon-McMillan-Breiman theorem [?] for every $\varepsilon > 0$ there exists a set $\Omega_\varepsilon \subset \Omega$ with measure $\ge 1 - \varepsilon$ and so that $\mu(A_{n-j}((x_1 \dots x_j)^\infty)) \ge e^{-(n-j)(h_\mu + \varepsilon)}$ for all n large enough and for all $(x_1 \dots x_j)^\infty$ so that $A_{n-j}((x_1 \dots x_j)^\infty) \cap \Omega_\varepsilon \ne \emptyset$. Hence

$$\begin{aligned}
\mu(\mathcal{C}_n(\delta)) &\ge e^{-\Delta^\beta} (1 - \psi^-(\Delta)) \sum_{A_n \in \mathcal{G}, T^j A_n \cap \Omega_\varepsilon \ne \emptyset} \mu(A_{j-\Delta}(T^j A_n)) e^{-(n-j)(h_\mu + \varepsilon)} \\
&\ge e^{-\Delta^\beta} (1 - \psi^-(\Delta)) e^{-(n-j)(h_\mu + \varepsilon)} \left(\sum_{x_1 \dots x_{j-\Delta}} \mu(A_{j-\Delta}(x_1 \dots x_{j-\Delta})) - \varepsilon\right) \\
&\ge e^{-\Delta^\beta - (n-j)(h_\mu + \varepsilon)} (1 - \psi^-(\Delta))\, (1 - \varepsilon)
\end{aligned}$$

and consequently $\lim_{n\to\infty} \frac{1}{n} \log \mu(\mathcal{C}_n(\delta)) \ge -(1 - \delta)(h_\mu + \varepsilon)$ if we take $\Delta = [n^\alpha]$ where $\alpha < 1$ is so that $\beta\alpha < 1$. Now let $\varepsilon \to 0^+$ in order to get the result. ∎

{LB}

Lemma 10 *Let* γ_μ *be as in Theorem 1. Then*

$$\liminf_{n\to\infty} \frac{1}{n} |\log \mu(\mathcal{C}_n(\delta))| \ge \gamma_\mu(1 - \delta).$$

The proof is exactly the same as in [4] Proposition 1(a). It uses the mixing properties of ψ-mixing measures without the assumption $\psi(0) < 1$.

4 Uniform decay rate of cylinders and the free energy

In this section we compute the free energy $F(\beta)$ of the process τ_n: it is defined by

$$F(\beta) \stackrel{def}{=} \lim_{n\to\infty} \frac{1}{n} \log \int_\Omega \exp(\beta\tau_n(A_n))\, d\mu = \lim_{n\to\infty} \frac{1}{n} \log \sum_{j=1}^{\infty} e^{\beta j} \mathbb{P}(\tau_n = j) \tag{5}$$ {FE}

whenever the limit exist.

{free.energy

Theorem 11 *Let μ be a weakly ψ-mixing measure. Then*

$$F(\beta) = \begin{cases} \beta & \text{if } -\gamma_\mu \le \beta < 0 \\ -\gamma_\mu & \text{if } \beta \le -\gamma_\mu \end{cases}.$$

Remark 1. Although $F(\beta)$ is not differentiable, one could still take its Legendre transform $\mathcal{L}F(\delta)$ and produce an upper bound for the deviation function $M(\delta)$ (see [19]). We immediately get

$$M(\delta) \le \mathcal{L}F(\delta) = -\gamma_\mu(1-\delta)$$

which is consistent with the bound obtained in Lemma 10.

Remark 2. The proof of the theorem splits into two parts. The first consists in getting an upper bound for the sum $\sum_{j=1}^{\infty} e^{\beta j}\mathbb{P}(\tau_n = j)$ which is achieved by using the mixing properties of the measure. We defer to the proof of this bound in [4] Proposition 6 which applies verbatim (it does not require the stringent condition $\psi(0) < 1$). This upper bound allows to show that the $\limsup_{n\to\infty} \frac{1}{n} \log \int_\Omega \exp(\beta\tau_n(A_n))\, d\mu$ is piecewise constant as prescribed in Theorem 11. However the lower bound is more interesting. Here we need an additional property of our measure, namely the existence of a sequence of cylinders whose measures decay exponentially to zero at a rate which is exactly the constant γ_μ given by Theorem 1 and whose first return is sublinear. This sequence is explicitly constructed in Lemma 13 below. We will give the proof of the lower bound after having proved Lemma 13.

As before let $\gamma_\mu = \liminf_n \frac{1}{n}|\log b_n|$ be the exponential decay rate of the measures of n-cylinders, where $b_n = \max_{A_n\in\mathcal{A}^n} \mu(A_n)$ and $0 < |\log\eta| \le \gamma_\mu \le h_\mu$ by Lemma 3.

{slowest.seq

Lemma 12 *There exists a sequence of n-cylinders A_n, $n = 1, 2, 3, \dots$, so that $\gamma_\mu = \lim_{n\to\infty} \frac{1}{n}|\log \mu(A_n)|$.*

Proof. We have to show that $\liminf$ is equal to the $\lim$ along a suitable sequence of cylinders. For this purpose let A_{n_j} (n_j increasing sequence) be a sequence of n_j-cylinders so that $\gamma_\mu = \lim_j \frac{1}{n_j}|\log\mu(A_{n_j})|$. Let $\varepsilon > 0$ and J large enough so that $\left|\frac{1}{n_j}|\log\mu(A_{n_j})| - \gamma_\mu\right| < \varepsilon/2$ for all $j \ge J$. Let $\alpha \in (0,1)$ (to be determined below), $\Delta = [n_j^\alpha]$ ($\Delta \ge \Delta_0$) and put $\tilde{A}_{n_j+(k-1)\Delta} = \bigcap_{i=0}^{k-1} T^{-i(n_j+\Delta)} A_{n_j}$ which implies by Lemma 2 that

$$\mu(\tilde{A}_{n_j+(k-1)\Delta}) = \mu(A_{n_j})^k \left(1 + \mathcal{O}(\psi^\pm(\Delta))\right)^{k-1}$$

As before for $\beta > 1$ we put

$$\mathcal{G}_k = \left\{A_{kn_j+(k-1)\Delta} \subset \tilde{A}_{kn_j+(k-1)\Delta} :\ \mu(A_{kn_j+(k-1)\Delta}) \ge e^{-k\Delta^\beta}\mu(\tilde{A}_{kn_j+(k-1)\Delta})\right\}.$$

Then if $|\mathcal{A}|^{(k-1)\Delta}e^{-k\Delta^\beta} < 1$ we get that $\exists$ at least one cylinder $A_{kn_j+(k-1)\Delta} \subset \tilde{A}_{kn_j+(k-1)\Delta}$, $A_{kn_j+(k-1)\Delta} \in \mathcal{A}^{kn_j+(k-1)\Delta}$, so that $\mu(A_{kn_j+(k-1)\Delta}) \ge e^{-k\Delta^\beta}\mu(\tilde{A}_{kn_j+(k-1)\Delta})$ and therefore

$$\begin{aligned}
\frac{\mu(A_{kn_j+(k-1)\Delta})}{kn_j+(k-1)\Delta} &\ge -\frac{k\Delta^\beta}{kn_j+(k-1)\Delta} + \frac{k\log\mu(A_{n_j})}{kn_j+(k-1)\Delta} + \frac{k\log(1-\psi^-(\Delta))}{kn_j+(k-1)\Delta} \\
&\ge -2\frac{\Delta^\beta}{n_j} + \frac{1}{1+\frac{\Delta}{n_j}}\frac{\log\mu(A_{n_j})}{n_j} - 2\frac{\psi^-(\Delta)}{n_j} \\
&\ge -cn_j^{\beta\alpha-1} + \frac{\log\mu(A_{n_j})}{n_j}
\end{aligned}$$

where we put $\Delta=[n_j^\alpha]$ and $c\approx 3+2\gamma_\mu$ (as $\frac{1}{1+\Delta/n_j}\le 1+2\frac{\Delta}{n_j}$ for j large enough) Hence

$$\left|\frac{|\log\mu(A_{kn_j+(k-1)\Delta})|}{kn_j+(k-1)\Delta}-\gamma_\mu\right|<\frac{\varepsilon}{2}+\frac{c}{n_j^{1-\beta\alpha}}<\varepsilon$$

for all k if n_j is large enough. Choose β so that $|\mathcal{A}|^\Delta e^{-\Delta^\beta}<1$ where $\Delta\sim n^\alpha$ ($\Delta\ge\Delta_0$). Then let $\alpha<1$ be so that $\alpha\beta<1$. ∎

{GC}

Lemma 13 *There exists a sequence of cylinders $B_j\in\mathcal{A}^j$ so that*

$$\lim_{j\to\infty}\frac{1}{j}|\log\mu(B_j)|=\gamma_\mu \quad \textit{and} \quad \lim_{j\to\infty}\frac{1}{j}\tau(B_j)=0$$

Proof. By Lemma 12 there exists a sequence of cylinders $A_n\in\mathcal{A}^n$ so that $\frac{1}{n}|\log\mu(A_n)|\to\gamma_\mu$ as $n\to\infty$. Let $\varepsilon>0$ and N so that $\left|\frac{1}{n}|\log\mu(A_n)|-\gamma_\mu\right|\le\varepsilon/3\ \forall\ n\ge N$. Let $\alpha,\alpha'\in(0,1)$ put $k_n=[n^{\alpha'}]$, $\Delta_n=[n^\alpha]$ and put for simplicity $n'=n+\Delta_n$, $(n+1)'=n+1+\Delta_{n+1}$. Then

$$k_{n+1}(n+1)'-\Delta_{n+1}-(k_nn'-\Delta_n)\in\left\{\begin{array}{ccc}[0,3] & \text{if} & k_{n+1}=k_n\\ {[k_n+\Delta_n,k_n+\Delta_n+3]} & \text{if} & k_{n+1}=k_n+1\end{array}\right..$$

Let $\epsilon_n=k_{n+1}-k_n$ ($\epsilon_n=0,1$) and for $j\in[k_nn'-\Delta_n,k_{n+1}(n+1)'-\Delta_{n+1})$ put

$$\mathcal{D}_j=\left\{D\in\mathcal{A}^{(k_n+\epsilon_n)n'-\Delta_n}:\ D\subset\tilde{A}_n\right\},$$

where

$$\tilde{D}_j=\bigcap_{j=0}^{k_n+\epsilon_n-1}T^{-jn'}A_n\in\bigvee_{j=0}^{k_n+\epsilon_n-1}T^{-jn'}\mathcal{A}^n.$$

For $\beta>1$ we define the 'good' set of cylinders in $\tilde{\mathcal{A}}^j$ whose measures are comparable to the measure of $\tilde{D}_n$.

$$\mathcal{G}_j=\left\{D\in\mathcal{A}^j:\ \mu(D)\ge e^{-(k_n+\epsilon_n)\Delta_n^\beta}\mu(\tilde{D}_j)\right\}$$

If $|\mathcal{A}|^\Delta e^{-\Delta^\beta}<1$ then $\mathcal{G}_j\ne\emptyset$. Hence we can find a j-cylinder $B_j\in\mathcal{G}_j$ so that $B_j\subset\tilde{D}_j$ and moreover has comparable measure: $\mu(B_j)\ge e^{-(k_n+\epsilon_n)\Delta_n^\beta}\mu(\tilde{D}_j)$. By the mixing property

$$\mu(\tilde{D}_j)=\left(1+\mathcal{O}(\psi^\pm(\Delta_n))\right)^{k_n+\epsilon_n-1}\mu(A_n)^{k_n+\epsilon_n}$$

which implies

$$\log\mu(B_j)\ge-(k_n+\epsilon_n)\Delta_n^\beta+k_n\log\mu(A_n)+k_n\log(1-\psi^-(\Delta_n))$$

If $\alpha'+\beta\alpha<1$ and n is large enough then $\frac{1}{j}(k_n+\epsilon_n)\Delta_n^\alpha<\frac{\varepsilon}{3}$ and $\frac{1}{j}k_n\log(1-\psi^-(\Delta_n))<\frac{\varepsilon}{3}$. Hence $\left|\frac{1}{j}|\log\mu(B_j)|-\gamma_\mu\right|<\varepsilon$ for all large enough j. Moreover we note that $\tau(A_j')\le n+\Delta_n$ which implies $\lim_{j\to\infty}\frac{1}{j}\tau(B_j)=0$. ∎

Proof of Theorem 11 As described in Remark 2 it will be sufficient to show that

$$\liminf_{n\to\infty}\frac{1}{n}\log\int_\Omega\exp(\beta\tau_n(A_n))d\mu\ge\left\{\begin{array}{cc}\beta & \text{if}\quad-\gamma_\mu\le\beta<0\\ -\gamma_\mu & \text{if}\quad\beta\le-\gamma_\mu\end{array}\right..$$

We have two cases.

(i): $-\gamma_\mu\le\beta<0$. The result immediately follows since

$$\sum_{j=1}^\infty\exp(\beta j)\mathbb{P}(\tau_n=j))\ge\exp(\beta n+\Delta)$$

where Δ was introduced at the beginning of Sect. 3.

(ii): $\beta < -\gamma_\mu$. Let us choose in any partition $\mathcal{A}^n$ a cylinder A'_n which verifies Lemma (13). Then

$$\sum_{j=1}^{\infty} e^{\beta j}\mathbb{P}(\tau_n = j) \geq \exp(\beta\tau_n(A'n))\mu(A'_n).$$

But $\mu(A'_n)$ decays exponentially fast to zero with a rate given by $-\gamma_\mu$, while $\frac{1}{n}\tau_n(A'n)$ goes to zero. This concludes the proof. ∎

References

[1] M Abadi, Sharp error terms and necessary conditions for exponential hitting times in mixing processes, *Ann. Probab.*, **32** (2004), 243–264.

[2] M Abadi, Hitting, returning and the short correlation function, *Bull. Braz. Math. Soc.*, **37** (2006), 593–609.

[3] M Abadi and A Galves, Inequalities for the occurrence times of rare events in mixing processes. The state of the art, *Markov Proc. Relat. Fields.*, **7** (2001), 97–112.

[4] M Abadi and S Vaienti: Large Deviations for Short Recurrence; to appear in *Disc. Cont. Dyn. Syst.*

[5] V Afraimovich, Pesin's dimension for Poincaré recurrence, *Chaos* **7**, 11-20, (1997).

[6] V Afraimovich, J-R Chazottes and B Saussol, Pointwise dimensions for Poincar recurrence associated with maps and special flows, *Disc. Cont. Dyn. Syst.*, **9** (2003), 263–280.

[7] V Afraimovich, E Ugalde and J Urias, "Fractal Dimensions for Poincaré Recurrence", Monograph Series on Nonlinear Sciences and Complexity, Vol. 2, Elsevier, (2006)

[8] L Barreira, Ya Pesin, J Schmeling, On a General Concept of Multifractality: Multifractal Spectra for Dimensions, Entropies and Lyapunov Exponents, *Chaos*, **7:1**, 27-38, (1997)

[9] C Beck and F Schlögl, "Thermodynamics of Chaotic Systems", Cambridge University Press, 1993.

[10] D Bessis, G Paladin, G Turchetti and S Vaienti, Generalized dimensions, entropies and Lyapunov exponents from the pressure function for strange sets, *J. Stat. Phys.*, **51** (1988), 109–134.

[11] P Billingsley, "Convergence in Probability Measures", Wiley, (1968)

[12] C Bonanno, S Galatolo and S Isola, Recurrence and algorithmic information, *Nonlinearity*, **17** (2004), 1057–1074.

[13] R Bowen, "Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms", Lecture Notes in Math 470, Springer-Verlag, New York, 1975.

[14] J-R Chazottes and F Durand, Local rates of Poincaré recurrence for rotations and weak-mixing, *Disc. Cont. Dyn. Syst.*, **12** (2005), 175–183.

[15] P Collet, A Galves, B Schmitt, Fluctuations of repetition times for Gibbsian sources, *Nonlinearity*, **12**, 1225-1237, (1999)

[16] A Csordas and P Szepfalusy, Generalized entropy decay rate of one-dimensional maps, *Phys. Rev. A*, **38** (1989), 2582–2587.

[17] P Doukhan, "Mixing Properties and Examples", Lecture Notes in Statistics 85, Springer-Verlag, 1995.

[18] J-P Eckmann and D Ruelle, Ergodic theory of chaos and strange attractors, *Rev. Mod. Phys.*, **57** (1985), 617–656.

[19] R S Ellis, “Entropy, Large Deviations and Statistical Mechanics”, Springer-Verlag, 1985.

[20] P Ferrero, N Haydn and S Vaienti, Entropy fluctuations for parabolic maps, *Nonlinearity*, **16**, (2003), 1203–1218

[21] A Galves and B Schmitt, Inequalities for hitting times in mixing dynamical systems, *Random Comput. Dyn.*, **5** (1997), 337–348.

[22] P Grassberger and I Procaccia, Estimation of the Kolmogorov entropy from a chaotic signal, *Phys. Rev. A*, **28**, 2591–2593, (1983)

[23] P Grassberger and I Procaccia, Dimensions and entropies of strange attractors from a fluctuating dynamics approach, *Phys. D*, **13** (1984), 34–54.

[24] N Haydn, Y Lacroix and S Vaienti: Hitting and Return Times in Ergodic Dynamical Systems: *Ann. of Probab.* **33** (2005), 2043–2050

[25] N Haydn and S Vaienti, The limiting distribution and error terms for return time of dynamical systems, *Disc. Cont. Dyn. Syst.*, **10** (2004), 584–616.

[26] M Hirata, B Saussol and S Vaienti, Statistics of return times: a general framework and new applications, *Comm. Math. Phys.*, **206** (1999), 33–55.

[27] M Kupsa, Local return rates in Sturmian Sub-shifts, *Acta Univ. Carol., Math. et Phys.*, **44** (2003), 17–28.

[28] P Kurka, Local return rates in substitutive sub-shifts, *Acta Univ. Carol., Math. et Phys.*, **44** (2003), 29–42.

[29] G Paladin, G Parisi and A Vulpiani, Intermittency in chaotic systems and Rényi entropies, *J. Phys. A*, **19** (1986), L997–L1001.

[30] V Penné, B Saussol and S Vaienti, Dimensions for recurrence times: topological and dynamical properties, *Disc. Cont. Dyn. Syst.*, **5** (1999), 783–798.

[31] Ya Pesin, “Dimension Theory in Dynamical Systems”, University of Chicago Press, (1997)

[32] A Rényi, “Probability Theory”, North Holland, Amsterdam, (1970)

[33] B Saussol, S Troubetzkoy and S Vaienti, Recurrence, dimensions and Lyapunov exponents, *J. Stat. Phys.*, **106** (2002), 623–634.

[34] F Takens and E Verbitsky: Generalized entropies: Rényi and correlation integral approach, *Nonlinearity*, **11** (1998), no. 4, 771–782

[35] F Takens and E Verbitsky: Multifractal analysis of local entropies for expansive homeomorphisms with specification, *Comm. Math. Phys.*, **203**, 593–612, (1999)